Population-adjusted treatment comparisons

Estimates based on Matching-Adjusted Indirect Comparisons (MAIC) and Simulated Treatment Comparisons (STC)

David M. Phillippo[1], A. E. Ades[1], Sofia Dias[1], Stephen Palmer,[2] Keith R. Abrams,[3] Nicky J. Welton[1]

[1]School of Social and Community Medicine, University of Bristol, [2]Centre for Health Economics, University of York,
[3]Department of Health Sciences, University of Leicester

Aims

- Review the properties and assumptions of methods for population-adjusted treatment comparison, including Matching-Adjusted Indirect Comparison (MAIC) and Simulated Treatment Comparison (STC).
- Provide guidance on their use in health technology appraisal (HTA).

More information is available in NICE DSU Technical Support Document 18 (Phillippo et al. 2016).

Background

In HTA submissions, a company wishes to compare their treatment *B* with that of a competitor, *C*. Standard indirect comparison and network meta-analysis assume that there are no differences in effect modifiers between the populations, and require a common comparator or a connected network — neither of which may be the case.

If effect modification is present on a given scale, relative effects $d_{u(P)} = g\left(Y_{u(P)}\right) - g\left(Y_{t(P)}\right)$ between treatments on that scale are specific to a given population *P*, where $Y_{t(P)}$ and $Y_{u(P)}$ are the mean outcomes on each treatment.

In an ideal scenario, individual patient data (IPD) would be available on all trials, and an IPD Network Meta-Regression could be performed. However, it is much more likely that a company only has access to IPD on their own trials and published aggregate summaries from their competitor's.

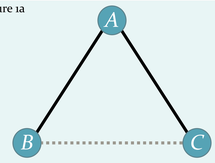
Population adjustment methods seek to use available IPD to adjust for any between-trial differences, or even reconcile unconnected networks, under certain constancy assumptions (Figure 1).

In a connected network with *AB* and *AC* trials, an anchored comparison can be made using randomisation with a common comparator *A* (Figure 1a).

In an unconnected network where there is no common comparator or there are single-arm studies, an unanchored comparison is the only option (Figure 1b).

Figure 1: Forms of indirect comparisons and constancy assumptions

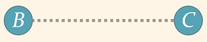
Figure 1a



Anchored

$$d_{BC(AC)} = d_{AC(AC)} - d_{AB(AC)}$$

Figure 1b



Unanchored

$$d_{BC(C)} = g\left(Y_{C(C)}\right) - g\left(Y_{B(C)}\right)$$

	Anchored		Unanchored	
Form of comparison:	Standard indirect comparison	Anchored population-adjusted indirect comparison	Naïve comparison, never used	Unanchored population-adjusted indirect comparison
Constancy assumption:	Constancy of relative effects $d_{AB(AB)} = d_{AB(AC)} = d_{AB}$	Conditional constancy of relative effects Predict $d_{AB(AC)}$ from *AB* trial	Constancy of absolute effects $Y_{B(B)} = Y_{B(C)} = Y_B$	Conditional constancy of absolute effects Predict $Y_{B(C)}$ from *B* trial
Valid only if:	No effect modifiers in imbalance	All effect modifiers known and adjusted for	No effect modifiers or prognostic variables in imbalance	All effect modifiers and prognostic variables known and adjusted for
Data:	Only requires aggregate data	Requires IPD on at least one trial	Only requires aggregate data	Requires IPD on at least one trial

Methods for population adjustment

Population adjustment methods are broadly of two types:

- Propensity score reweighting, such as Matching-Adjusted Indirect Comparison (MAIC; Signorovitch et al. 2010), where individuals in the *AB* trial are weighted so that the reweighted covariate distribution matches that of the aggregate *AC* trial.
- Outcome regression, such as Simulated Treatment Comparison (STC; Caro and Ishak 2010), where a model is fitted in the *AB* trial and used to predict outcomes in the aggregate *AC* trial.

Target population and shared effect modifier assumption

The results of a population-adjusted analysis are irrelevant if they cannot be obtained for the correct target population. The shared effect modifier assumption holds for active treatments *B* and *C* if:

a) *B* and *C* have the same effect modifiers, and

b) The change in treatment effect caused by each effect modifier is the same for *B* and *C*

If this is the case, then the relative effect d_{BC} is valid for *any* population. The shared effect modifier assumption is evaluated on a clinical basis, and is more likely to be satisfied by treatments in the same class.

Recommendations

The focus of the following recommendations is statistical and clinical validity, transparency, and consistency in the use of population adjustment methods for health technology appraisal.

Recommendation 1 When connected evidence with a common comparator is available, a population-adjusted anchored indirect comparison may be considered. Unanchored indirect comparisons may only be considered in the absence of a connected network of randomised evidence, or where there are single-arm studies involved.

Unanchored comparisons require much stronger assumptions, so anchored comparisons are always preferred.

Recommendation 2 Submissions using population-adjusted analyses in a connected network need to provide evidence that they are likely to produce less biased estimates of treatment differences than could be achieved through standard methods.

a) Evidence must be presented that there are grounds for considering one or more variables as effect modifiers on the appropriate transformed scale. This can be empirical evidence, or an argument based on biological plausibility.

b) Quantitative evidence must be presented that population adjustment would have a material impact on relative effect estimates due to the removal of substantial bias.

Justification is required for moving away from standard anchored methods. This is in line with the NICE Methods Guide, which states that "treatment effect modifiers should be identified before data analysis, either by a thorough review of the subject area or discussion with experts in the clinical discipline."

Recommendation 3 Submissions using population-adjusted analyses in an unconnected network need to provide evidence that absolute outcomes can be predicted with sufficient accuracy in relation to the relative treatment effects, and present an estimate of the likely range of residual systematic error in the "adjusted" unanchored comparison.

If this evidence cannot be provided, the amount of bias in an unanchored comparison is unknown and likely to be substantial.

Recommendation 4 The following variables should be adjusted for in a population-adjusted analysis:

a) For an anchored indirect comparison, propensity score weighting methods should adjust for all effect modifiers (in imbalance or not), but no prognostic variables. Outcome regression methods should adjust for all effect modifiers in imbalance, and any other prognostic variables and effect modifiers that improve model fit.

b) For an unanchored indirect comparison, both propensity score weighting and outcome regression methods should adjust for all effect modifiers and prognostic variables, in order to reliably predict absolute outcomes.

For anchored comparisons, only adjusting for effect modifiers minimises bias without unnecessarily reducing precision. Unanchored comparisons require all covariates to be adjusted for, as predictions of absolute outcomes are required.

Recommendation 5 Indirect comparisons should be carried out on the linear predictor scale, with the same link functions that are usually employed for those outcomes.

In line with general modelling practice (NICE, ISPOR). Effect modification is defined with respect to this scale, so must also be clinically meaningful.

Recommendation 6 The target population for any treatment comparison must be explicitly stated, and population-adjusted estimates of the relative treatment effects must be generated for this target population.

Population adjustment methods are only useful for decision making if they can produce estimates for the appropriate target population; the shared effect modifier assumption may be utilised if appropriate.

Processes for population-adjusted indirect comparison

Anchored

Propensity score reweighting	Outcome regression
1. Provide evidence for effect modifier status on a suitable transformed scale.	
2. Provide evidence that effect modifiers are in substantial imbalance between studies.	
3a. Create a logistic propensity score model, which includes all effect modifiers but no prognostic variables. This is equivalent to a model on the log of the weights: $\log\left(w_{it}\right) = \alpha_0 + \boldsymbol{\alpha}_1^T \boldsymbol{X}_{it}^{EM}$	**3.** Fit an outcome model in the *AB* trial, which includes all effect modifiers in imbalance and any other prognostic variables or effect modifiers that improve model fit: $g\left(\mu_{t(AB)}\left(\boldsymbol{X}\right)\right) = \beta_0 + \boldsymbol{\beta}_1^T \boldsymbol{X} + \left(\beta_B + \boldsymbol{\beta}_2^T \boldsymbol{X}^{EM}\right) I\left(t = B\right)$
3b. Estimate the weights using the method of moments to match effect modifier distributions between trials. This is equivalent to minimising $\sum_{t=A,B} \sum_{i=1}^{N_{t(AB)}} \exp\left(\boldsymbol{\alpha}_1^T \boldsymbol{X}_{it}^{EM}\right)$ when $\bar{\boldsymbol{X}}_{(AC)}^{EM} = \boldsymbol{0}$.	
4. Predict outcomes on treatments *A* and *B* in the *AC* trial by reweighting the outcomes of the *AB* individuals: $\hat{Y}_{t(AC)} = \frac{\sum_{i=1}^{N_{t(AB)}} Y_{t(AB)} \hat{w}_{it}}{\sum_{i=1}^{N_{t(AB)}} \hat{w}_{it}}$	**4.** Predict transformed outcomes on treatments *A* and *B* in the *AC* trial using the outcome model: $g\left(\hat{Y}_{t(AC)}\right) = \hat{\beta}_0 + \hat{\boldsymbol{\beta}}_1^T \bar{\boldsymbol{X}}_{(AC)} + \left(\hat{\beta}_B + \hat{\boldsymbol{\beta}}_2^T \bar{\boldsymbol{X}}_{(AC)}^{EM}\right) I\left(t = B\right)$
5. Form the anchored indirect comparison in the *AC* population as: $\hat{\Delta}_{BC(AC)} = g\left(\bar{Y}_{C(AC)}\right) - g\left(\bar{Y}_{A(AC)}\right) - \left(g\left(\hat{Y}_{B(AC)}\right) - g\left(\hat{Y}_{A(AC)}\right)\right)$	
6. Calculate standard errors using a robust sandwich estimator, bootstrapping, or Bayesian techniques.	**6.** Calculate standard errors using the outcome model.
7. If justified, use the shared effect modifier assumption to transport the $\hat{\Delta}_{BC(AC)}$ estimate into the target population for the decision. Otherwise, comment on the representativeness of the *AC* population to the true target population.	
8. Present the distribution of estimated weights, and effective sample size.	**8.** Present standard model fit statistics.

Unanchored

Propensity score reweighting	Outcome regression
1a. Create a logistic propensity score model, which includes all effect modifiers and prognostic variables. This is equivalent to a model on the log of the weights: $\log\left(w_i\right) = \alpha_0 + \boldsymbol{\alpha}_1^T \boldsymbol{X}_i$	**1.** Fit an outcome model in the *A* trial, which includes all effect modifiers and prognostic variables: $g\left(\mu_{B(B)}\left(\boldsymbol{X}\right)\right) = \beta_0 + \boldsymbol{\beta}_1^T \boldsymbol{X} + \left(\beta_B + \boldsymbol{\beta}_2^T \boldsymbol{X}^{EM}\right)$
1b. Estimate the weights using the method of moments to match effect modifier distributions between trials. This is equivalent to minimising $\sum_{i=1}^{N_{B(B)}} \exp\left(\boldsymbol{\alpha}_1^T \boldsymbol{X}_i\right)$ when $\bar{\boldsymbol{X}}_{(C)}^{EM} = \boldsymbol{0}$.	
2. Predict outcomes on treatment *B* in the *C* trial by reweighting the outcomes of the *B* individuals: $\hat{Y}_{B(C)} = \frac{\sum_{i=1}^{N_{B(B)}} Y_{B(B)} \hat{w}_i}{\sum_{i=1}^{N_{B(B)}} \hat{w}_i}$	**2.** Predict transformed outcomes on treatments *A* and *B* in the *C* trial using the outcome model: $g\left(\hat{Y}_{B(C)}\right) = \hat{\beta}_0 + \hat{\boldsymbol{\beta}}_1^T \bar{\boldsymbol{X}}_{(C)} + \left(\hat{\beta}_B + \hat{\boldsymbol{\beta}}_2^T \bar{\boldsymbol{X}}_{(C)}^{EM}\right)$
3. Form the unanchored indirect comparison in the *C* population as: $\hat{\Delta}_{BC(C)} = g\left(\bar{Y}_{C(C)}\right) - g\left(\hat{Y}_{B(C)}\right)$	
4. Calculate standard errors using a robust sandwich estimator, bootstrapping, or Bayesian techniques.	**4.** Calculate standard errors using the outcome model.
5. Provide evidence that absolute outcomes can be predicted with sufficient accuracy in relation to the relative treatment effects, and present an estimate of the likely range of residual systematic error. If this evidence cannot be provided or is limited, then state that the amount of bias in the indirect comparison is likely to be substantial, and could even exceed the magnitude of treatment effects which are being estimated.	
6. If justified, use the shared effect modifier assumption to transport the $\hat{\Delta}_{BC(C)}$ estimate into the target population for the decision. Otherwise, comment on the representativeness of the *C* population to the true target population.	
7. Present the distribution of estimated weights, and effective sample size.	**7.** Present standard model fit statistics.

References

Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. 2016. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Available from www.nicedsu.org.uk/.

Caro JJ, Ishak KJ. 2010. No head-to-head trial? Simulate the missing arms. Pharmacoeconomics. 28(10):957-967.

Signorovitch JE, Wu EQ, Yu AP, Gerrits CM, Kantor E, Bao YJ, Gupta SR, Mulani PM. 2010. Comparative effectiveness without head-to-head trials a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. Pharmacoeconomics. 28(10):935-945.

@dmphillippo
david.phillippo@bristol.ac.uk

NICE The TSD was commissioned and funded by the Decision Support Unit at the National Institute for Health and Care Excellence.

MRC Medical Research Council Ongoing work is supported by the Medical Research Council grant number MR/P015298/1.

University of BRISTOL